SEQUENCE LISTING

<110> XIANG, JINHUA
WÜNSCHMANN, SABINA
SCHMIDT, WARREN
STAPLETON, JACK T.

<120> FULL-LENGTH GB VIRUS C (HEPATITIS G VIRUS) RNA TRANSCRIPTS ARE INFECTIOUS IN PRIMARY CD4 POSITIVE T CELLS

<130> IOWA:030US

<140> UNKNOWN
<141> 2001-04-05

<140> 60/253,390
<141> 2000-11-27

<150> 60/195,597
<151> 2000-04-06

<160> 2

<170> PatentIn Ver. 2.1

<210> 1
<211> 9395
<212> DNA
<213> Hepatitis G virus

<400> 1
tgacgtgggg gggttgatcc ccccccccg gcactgggtg caagccccat aaaccgacgc 60
ctatctaagt agacgcaatg actcggcgcc gactcggcga ccggccaaaa ggtggtggat 120
gggtggtgac agggttggta ggtcgtaaat cccggtcatc ctggtagcca ctataggtgg 180
gtcttaagag aaggtcaaga ctcctcttgt gcctgcggcg agaccgcgca cggtccacag 240
gtgctggccc taccggtgtg aataagggcc cgacgtcagg ctcgtcgtta aaccgagccc 300
gtcacccacc tgggcaaacg acgcccacgt acggtccacg tcgcccttca atgtctctct 360
tgaccaatag gtttatccgg cgagttgaca aggaccagtg ggggccgggg gttatgggga 420
aggaccccaa accctgccct tcccggtggg ccgggaaatg catggggcca cccagctccg 480
cggcggcctg cagccggggt agcccaagaa tccttcgggt gagggcgggt ggcatttctc 540
ttttctatac catcatggca gtccttctgc tccttctcgt ggttgaggcc ggggccattc 600
tggccccggc cacccacgct tgtcgagcga atgggcaata tttcctcaca aattgctgtg 660
ccccggaaga catcgggttc tgcctggaag gcggatgcct ggtggccctg gggtgcacgg 720
tttgcaccga ccgttgctgg ccactgtatc aggcgggttt ggctgtgcgg cctggcaagt 780
ccgcggccca gctcgttggg gaactgggga gcctgtacgg gcccttgtcg gtctcggctt 840
acgtagccgg gatcctgggt ctgggcgagg tttactccgg ggtcctgaca gttggtgttg 900
cgttgaggcg ccgggtctac ctgatgccca acctgaagtg tgcagtagaa tgtgacgtta 960
agtggggaag tgagttttgg agatggactg agcagttggc ctccaattac tggattttgg 1020
aataccttg gaaagtccca tttgaatttt ggagaggagt gatgagcctg acccctctgt 1080
tggtttgggt ggccgcattg cttttgctgg agcaacggat tgtcatggtt ttcctgctgg 1140
tgacgatggc ggggatgttg caaggcgccc ccgcctccgt tttggggtcc cgccccttg 1200
actacgggtt gaagtggcag tcatgctcct gcagggctaa cgggtcgcgt attcccactg 1260
gggagagggt gtgggatcga gggaatgtca cgctcttgtg tgactgcccc aacggccct 1320
gggtttgggt cccggccttt tgccaggcgg ttgggtgggg cgaccccatc acccattgga 1380
gccacggaca aaaccagtgg cccctatcat gcccccaata tgtctatggg tctgtgtccg 1440
taacgtgcgt gtggggttcc gtgtcttggt ttgcctcgac cggcggtcgt gattcgaaga 1500
tcgatgtgtg gagtttggtg ccggttggat ctgccagctg caccatagcc gctctagggt 1560

catcggatcg cgacacggtg gttgagctct ccgagtgggg agtcccgtgc gtaacgtgta 1620
ttctggaccg tcggcctgct tcatgtggca cctgtgtgcg ggactgctgg cccgaaaccg 1680
ggtcggttag attccctttc catcggtgcg gcacggggcc tcggctgaca aaggacttgg 1740
aagctgtgcc cttcgtcaac aggacaactc ccttcaccat aaggggcccc ctgggcaacc 1800
aggggagagg caacccggtg cggtcgcccc tgggttttgg gtcctacacc atgaccaaga 1860
tccgggattc cctgcatttg gtgaaatgtc ccacaccagc catagagcct ccgactggaa 1920
cgttcgggtt cttccccgga gtcccgccca ttaacaactg catgccgcta ggcacggaag 1980
tgtctgaggc attgggcgga gctgggctta cgggggggtt ctacgagcct ctggttcgca 2040
ggtgttcgga gctgatggga cgccgaaatc cggtttgccc ggggtacgca tggctgtcct 2100
ctggtagacc tgacgggttc atacacgtcc aggggcacct gcaggaggtg gatgcgggca 2160
acttcatccc tcctccacgc tggttgctct tggattttgt atttgtcctg ctctatctga 2220
tgaagctggc tgaggcacgg ttggtcccgt tgatcttgct tctgctgtgg tggtgggtga 2280
accagttggc ggttctagga ctgccggctg tggacgctgc cgtggcgggt gaagtttttg 2340
cgggccctgc cttgtcatgg tgtttgggcc ttcccactgt cagtatgata ctaggtctag 2400
caaacctggt gttgtacttt cggtggatgg gccctcagcg cctcatgttc ctcgtgttgt 2460
ggaagctcgc tcggggagct ttcccgctgg cacttttgat ggggatttcg gcgacccgcg 2520
ggcgcacctc tgtgctcggg gccgagttct gcttcgatgt cacattcgag gtggacactt 2580
cggtgttggg ctgggtggtg gccagcgtgg tggcttgggc catagcgctc ctgagctcaa 2640
tgagcgcagg ggggtggaag cacaaggccg tgatctatag gacgtggtgt aaagggtacc 2700
aggctgtgcg ccagagggtg gtgcggagcc ccctcgggga ggggcgtcct accaagcttc 2760
tgacgttcgc ctggtgcttg gcctcatata tctggccgga tgctgtgatg atggtggtgg 2820
tggccttggt cctcctcttc ggcctgttcg acgcactgga ctgggccctg gaggagctcc 2880
tggtctcccg gccctcgtta cggcgactgg cacgggtggt tgagtgctgt gtgatggcgg 2940
gcgagaaggc caccaccatc cgactggtct ccaagatgtg cgcaagaggg gcctacctgt 3000
ttgaccacat gggctctttc tcgcgcgctg tcaaggagcg cttgttggaa tgggacgcgg 3060
ctttggagcc cttgtcattc actaggacgg actgtcgcat catcagagat gccgcgagga 3120
ccctgtcctg cggacagtgc gtcatgggtt tacccgtggt agcacggcgc ggtgatgagg 3180
ttctcatcgg cgtctttcag gatgtgaatc atttgcctcc cgggtttgtc ccgactgcac 3240
cagttgtcat ccgtcggtgc ggaaagggct tcctgggggt cacgaaggca gccttgacag 3300
gtagggatcc tgacttacat ccagggaacg tcatggtgtt ggggacggct acgtcacgaa 3360
gcatgggcac atgtctgaat ggcctgctgt tcacaacttt ccatggggct tcatcccgaa 3420
ccatcgccac gcccgtgggg gcccttaatc ccaggtggtg gtcagccagt gatgacgtca 3480
cggtgtaccc gcttccagat ggggcaactt cgttgacgcc ctgcacttgc caggcggagt 3540
cctgttgggt tattagatcc gacggggctt tgtgccatgg cttgagcaag ggggacaagg 3600
ttgagctgga tgtggccatg gaggtctctg acttccgtgg ttcgtctggt tcaccggtcc 3660
tttgcgacaa agggcacgca gtaagaatgc tcgtgtcagt gctccactct ggcggcaggg 3720
ttactgcggc gcgattcact aggccgtgga ctcaagtacc aacagatgcc aagactacca 3780
cagaaccccc tccggtgccg gcaaaaggag ttttcaagga ggccccgttg tttatgccta 3840
cgggggcggg aaagagcacc cgcgtaccgt tggagtacgg caacatgggc cacaaggtct 3900
tgatcttgaa cccgtcggta gctaccgtga gggccatggg cccatacatg gagcggctgg 3960
cggggaaaca ccccagtatt tactgtggcc atgacaccac tgctttcaca aggatcactg 4020
actcgccccc tacgtattcc acttacggaa ggtttttggc caaccctagg cagatgctga 4080
ggggtgtgtc ggtggtcatt tgtgacgagt gccacagtca tgactcaact gtgttgttgg 4140
gcattgggcg tgtcagggag ctggcgcgag gatgtggagt gcaattggtg ctctacgcca 4200
ctgccacccc tcccggatcc ccgatgaccc agcacccatc aatcattgag acaaaactgg 4260
acgtgggaga gatccccttc tatgggcatg gcatacctct tgagcggatg cggaccggaa 4320
ggcatctcgt attctgccac tccaaggctg agtgcgagcg cctggcgggc cagttttcgg 4380
ctaggggggt aaatgccatc gcctattaca gggggaaaga cagttctatc atcaaagatg 4440
gagacctggt ggtgtgtgct acagacgcac tatccactgg gtacactggg aacttcgatt 4500
ctgtcaccga ttgtgggtta gtggtggagg aggtcgtcga ggtgaccctt gatcccacca 4560
ttaccatctc cctgcgcacg gtgcccgcgt cggctgaact gtcgatgcag cggcgaggac 4620
gcacgggtag gggcaggtct gggcgctact actacgcggg ggtcggcaag gcccctgctg 4680
gtgtggtgcg ctcaggtcct gtctggtcgg cggtggaagc cggtgtgacc tggtacggaa 4740
tggaacctga cctgacagca aacctactga gactttacga caactgccct tacaccgcag 4800
ccgtcgcagc tgacattggg gaagccgcgg tgttcttttc ggggcttgcc ccgttgagga 4860
tgcatcccga tgttagctgg gcaaaagttc gcggcgtcaa ctggcccttc ctggtgggtg 4920
ttcagcggac catgtgccgg gaaacactgt ctcccggccc atcggatgac ccccagtggg 4980

caggtctgaa gggcccgaat cctgtcccac tcctgctgag gtggggcaat gatttaccat 5040
ctaaagtggc cggccatcac atcgtggacg acctggtccg taggctcggg gtggcggagg 5100
gttacgtccg ctgcgatgcg ggacccatct tgatggtggg cctcgctatt gcgggggca 5160
tgatctatgc gtcatacacc gggtctctcg tggtggttac agactgggat gtgaaggggg 5220
gtggcagccc cctttatcgg catggagacc aggccacgcc ccagccggtt gtgcaggtcc 5280
ccccggtaga ccatcggccg gggggagagt ctgcgccatc ggatgccaac acagtgacag 5340
atgcggtggc ggccatccag gtggattgcg attggtcagt catgaccctg tcgatcgggg 5400
aagtgctgtc cttggcccag gctaagacgg ccgaggccta cgcagctacc accaagtggc 5460
ttgctggctg ctacacgggg acgcgggccg tccccactgt ttcaattgtt gacaagctct 5520
tcgccggggg ctgggcggcg gtggtaggcc attgccacag tgtaatagct gcggcagtgg 5580
cggcctatgg ggcttctagg agccctccat tggctgctgc cgcttcctac ctcatggggt 5640
tgggcgtcgg aggcaacgcg caaacccgct tagcctccgc tctcctacta ggggccgctg 5700
ggaccgctct gggcacgcct gtcgtggggt taaccatggc gggcgcgttc atgggaagtg 5760
ctagcgtctc cccctccttg gtcaccattt tactgggggc cgtgggggggc tgggagggcg 5820
tggtgaatgc ggctagcctt gtcttcgact ttatggcggg gaaactatca tcagaagatc 5880
tgtggtatgc catcccagtg ctaaccagtc cgggggcagg acttgcgggg atcgccctcg 5940
ggttggtgtt gtactcagct aacaactctg gcactaccac ttggttgaac cgtctgctga 6000
ctacattgcc aaggtcctca tgcatccctg acagttactt tcagcaggcc gattactgtg 6060
acaaggtctc agctgtgctc cgacgcttga gcctcactcg caccgtggtt gccctggtca 6120
acagggagcc taaggtggat gaggttcagg tggggtacgt ctgggacttg tgggagtgga 6180
tcatgcgtca agtgcgcatg gtgatggcca gacttcgggc cctctgcccc gtggtgtcat 6240
taccccttatg gcactgcggg gaggggtggt ccggagaatg gttgttggac ggccatgttg 6300
agagtcgttg tctttgtggt tgcgtgatca ccggtgatgt tttgaatggg caactcaaag 6360
atccagttta ctctaccaag ctgtgcaggc attattggat ggggacagtc cctgtgaaca 6420
tgctgggcta tggcgagacg tcgcctttgc tcgcctcaga caccccgaag gtggtaccat 6480
tcgggacgtc tgggtgggct gaggtggtgg tgacccctac ccacgttgtg atcaggcgaa 6540
catccgccta caaactgctg cgccagcaaa tcctgtcggc tgctgttgct gagccctatt 6600
acgtcgacgg cataccggtc tcatgggacg cggacgcgcg agcgcctgcc atggtctatg 6660
gccctgggca aagtgtcacc attgacgggg aacgctacac ccttccgcat caactgcggc 6720
ttaggaatgt ggcgccctct gaggtgtcat ccgaggtgtc cattgacatt gggacggaga 6780
ctgaagactc agaactgact gaggccgacc tgccgccggc ggctgcagcc cttcaggcta 6840
tcgagaatgc tgcgagaatt cttgaacctc acatagatgt catcatggaa gattgcagta 6900
caccctctct ttgtgggagt agccgagaga tgcctgtgtg gggagaagac ataccccgca 6960
ctccatcgcc agcacttatc tcggttactg agagcagccc agatgagaag accccgtcgg 7020
tgtcttcctc gcaggaggat accccgtctt ctgactcatt cgaggtcatc caagagtccg 7080
agacagccga aggggaggaa agcgtcttca acgtggctct ttccgtacta aaagccttgt 7140
ttccacagag cgatgccaca agaaagctta ccgttaagat gtcatgctgt gttgagaaga 7200
gcgtaacacg cttcttttca ttgggattga cggtcgctga cgtggcaagc ctgtgtgaga 7260
tggaaatcca gaaccataca gcctattgtg acaaggtgcg cactccgctt gaattgcagg 7320
ttgggtgctt ggtgggcaat gaacttacct ttgaatgtga caagtgtgag gctaggcaag 7380
agaccttggc ttccttctct tacatttggt ctggggtgcc actgacgagg gccactccgg 7440
ccaagccccc tgtggtgagg ccggttggct ccttgctggt ggccgacacc accaaggtgt 7500
atgtcaccaa cccggacaat gttgggagaa gagttgacaa ggttaccttc tggcgtgccc 7560
ctagggttca tgacaaattc ctcgtggact ccatagagcg cgctaagagg gcagctcaag 7620
cctgcctaag catgggttac acttatgagg aggcaataag gactgtaagg ccacatgctg 7680
ccatgggctg gggatctaag gtgtcggtca aggacctcgc caccctgcg gggaagatgg 7740
ctgtccatga ccggctccag gagatacttg aagggacgcc agtccccttt actcttactg 7800
tgaaaaagga agtgttcttc aaagaccgaa aggaagagaa ggcccccgc ctcattgtgt 7860
tccccccct ggacttccgg atagctgaaa agcttattct gggagaccct ggacgggtag 7920
ccaaggcggt gttgggggg gcctacgcct tccagtacac cccaaatcag cgaattaggg 7980
agatgctcaa actgtgggaa tcaaagaaga caccatgcgc catctgtgtg gacgccacat 8040
gcttcgacag tagcataact gaagaggacg tggcgctgga gacagagctt tatgccctgg 8100
cttcagacca tccagaatgg gtgcgtgccc tggggaaata ctatgcctct ggcacaatgg 8160
taacccccga gggggtgcca gtgggtgaga ggtattgtag atcctcaggg gtcttgacca 8220
ccagtgcgag caactgcttg acttgctata tcaaggtgaa agccgcctgt gagagggtgg 8280
ggctgaaaaa tgtctcgctc ctcatcgctg gcgatgactg tttgatcata tgcgaacggc 8340
ctgtgtgcga tcctagcgac gctttgggca gagccctggc gagctacggg tacgcatgcg 8400

```
agccttcgta tcatgcatca ctggacacgg ccccttctg ctccacttgg ctagctgagt 8460
gcaatgcaga tgggaaacgc catttcttcc tgaccacgga ctttcggagg cccctcgctc 8520
gcatgtcgag cgagtacagt gacccaatgg cttcggccat cggttacatc ctcctatacc 8580
cttggcatcc tatcacacgg tgggtcatca tccctcacgt gctcacctgc gcgtttaggg 8640
gtggtggcac accgtctgat cctgtgtggt gccaggtaca tggtaattac tacaagtttc 8700
cactggacaa actgcctaac atcatcgtgg ccctccacgg accagcagcg ttgagggtta 8760
ccgcagacac aactaagaca aaaatggagg ctggcaaggt gctgagcgac ctcaagctcc 8820
ctggcctagc agtccaccgg aagaaggccg gggcattgcg aacgcgtatg ctccggtcgc 8880
gcggttgggc tgagttggct agggggctgt tgtggcgtcc aggcctgcgg cttccccctc 8940
cggagattgc tggtatcccc gggggtttcc ccctttcccc cccctatatg ggggtggttc 9000
atcaattgga tttcacaagc cagaggagtc gctggcggtg gttggggttc ttagccctgc 9060
tcatcgtagc cctcttcggg tgaactaaat tcatctgttg cggcaaggtc cggtgactga 9120
tcatcactgg aggaggttcc cgccctcccc gccccagggg tctccccgct gggtaaaaag 9180
ggcccggcct tgggaggcat ggtggttact aaccccctgg cagggtcaaa gcctgatggt 9240
gctaatgcac tgccacttcg gtggcgggtc gctaccttat agcgtaatcc gtgactacgg 9300
gctgctcgca gagccctccc cggatggggc acagtgcact gtgatctgaa ggggtgcacc 9360
ccggtaagag ctcggcccaa aggccgggtt ctact                            9395
```

<210> 2
<211> 2910
<212> PRT
<213> Hepatitis G virus

<400> 2

```
Met Ser Leu Leu Thr Asn Arg Phe Ile Arg Arg Val Asp Lys Asp Gln
  1               5                  10                  15

Trp Gly Pro Gly Val Met Gly Lys Asp Pro Lys Pro Cys Pro Ser Arg
             20                  25                  30

Trp Ala Gly Lys Cys Met Gly Pro Pro Ser Ser Ala Ala Ala Cys Ser
         35                  40                  45

Arg Gly Ser Pro Arg Ile Leu Arg Val Arg Ala Gly Gly Ile Ser Leu
     50                  55                  60

Phe Tyr Thr Ile Met Ala Val Leu Leu Leu Leu Leu Val Val Glu Ala
 65                  70                  75                  80

Gly Ala Ile Leu Ala Pro Ala Thr His Ala Cys Arg Ala Asn Gly Gln
                 85                  90                  95

Tyr Phe Leu Thr Asn Cys Cys Ala Pro Glu Asp Ile Gly Phe Cys Leu
            100                 105                 110

Glu Gly Gly Cys Leu Val Ala Leu Gly Cys Thr Val Cys Thr Asp Arg
        115                 120                 125

Cys Trp Pro Leu Tyr Gln Ala Gly Leu Ala Val Arg Pro Gly Lys Ser
    130                 135                 140

Ala Ala Gln Leu Val Gly Glu Leu Gly Ser Leu Tyr Gly Pro Leu Ser
145                 150                 155                 160

Val Ser Ala Tyr Val Ala Gly Ile Leu Gly Leu Gly Glu Val Tyr Ser
                165                 170                 175
```

```
Gly Val Leu Thr Val Gly Val Ala Leu Arg Arg Arg Val Tyr Leu Met
            180                 185                 190

Pro Asn Leu Lys Cys Ala Val Glu Cys Asp Val Lys Trp Gly Ser Glu
        195                 200                 205

Phe Trp Arg Trp Thr Glu Gln Leu Ala Ser Asn Tyr Trp Ile Leu Glu
    210                 215                 220

Tyr Leu Trp Lys Val Pro Phe Glu Phe Trp Arg Gly Val Met Ser Leu
225                 230                 235                 240

Thr Pro Leu Leu Val Trp Val Ala Ala Leu Leu Leu Leu Glu Gln Arg
                245                 250                 255

Ile Val Met Val Phe Leu Leu Val Thr Met Ala Gly Met Leu Gln Gly
            260                 265                 270

Ala Pro Ala Ser Val Leu Gly Ser Arg Pro Phe Asp Tyr Gly Leu Lys
        275                 280                 285

Trp Gln Ser Cys Ser Cys Arg Ala Asn Gly Ser Arg Ile Pro Thr Gly
    290                 295                 300

Glu Arg Val Trp Asp Arg Gly Asn Val Thr Leu Leu Cys Asp Cys Pro
305                 310                 315                 320

Asn Gly Pro Trp Val Trp Val Pro Ala Phe Cys Gln Ala Val Gly Trp
                325                 330                 335

Gly Asp Pro Ile Thr His Trp Ser His Gly Gln Asn Gln Trp Pro Leu
            340                 345                 350

Ser Cys Pro Gln Tyr Val Tyr Gly Ser Val Ser Val Thr Cys Val Trp
        355                 360                 365

Gly Ser Val Ser Trp Phe Ala Ser Thr Gly Gly Arg Asp Ser Lys Ile
    370                 375                 380

Asp Val Trp Ser Leu Val Pro Val Gly Ser Ala Ser Cys Thr Ile Ala
385                 390                 395                 400

Ala Leu Gly Ser Ser Asp Arg Asp Thr Val Val Glu Leu Ser Glu Trp
                405                 410                 415

Gly Val Pro Cys Val Thr Cys Ile Leu Asp Arg Arg Pro Ala Ser Cys
            420                 425                 430

Gly Thr Cys Val Arg Asp Cys Trp Pro Glu Thr Gly Ser Val Arg Phe
        435                 440                 445

Pro Phe His Arg Cys Gly Thr Gly Pro Arg Leu Thr Lys Asp Leu Glu
    450                 455                 460

Ala Val Pro Phe Val Asn Arg Thr Thr Pro Phe Thr Ile Arg Gly Pro
465                 470                 475                 480
```

```
Leu Gly Asn Gln Gly Arg Gly Asn Pro Val Arg Ser Pro Leu Gly Phe
                485                 490                 495

Gly Ser Tyr Thr Met Thr Lys Ile Arg Asp Ser Leu His Leu Val Lys
            500                 505                 510

Cys Pro Thr Pro Ala Ile Glu Pro Pro Thr Gly Thr Phe Gly Phe Phe
        515                 520                 525

Pro Gly Val Pro Pro Ile Asn Asn Cys Met Pro Leu Gly Thr Glu Val
    530                 535                 540

Ser Glu Ala Leu Gly Gly Ala Gly Leu Thr Gly Gly Phe Tyr Glu Pro
545                 550                 555                 560

Leu Val Arg Arg Cys Ser Glu Leu Met Gly Arg Arg Asn Pro Val Cys
                565                 570                 575

Pro Gly Tyr Ala Trp Leu Ser Ser Gly Arg Pro Asp Gly Phe Ile His
            580                 585                 590

Val Gln Gly His Leu Gln Glu Val Asp Ala Gly Asn Phe Ile Pro Pro
        595                 600                 605

Pro Arg Trp Leu Leu Leu Asp Phe Val Phe Val Leu Leu Tyr Leu Met
    610                 615                 620

Lys Leu Ala Glu Ala Arg Leu Val Pro Leu Ile Leu Leu Leu Leu Trp
625                 630                 635                 640

Trp Trp Val Asn Gln Leu Ala Val Leu Gly Leu Pro Ala Val Asp Ala
                645                 650                 655

Ala Val Ala Gly Glu Val Phe Ala Gly Pro Ala Leu Ser Trp Cys Leu
            660                 665                 670

Gly Leu Pro Thr Val Ser Met Ile Leu Gly Leu Ala Asn Leu Val Leu
        675                 680                 685

Tyr Phe Arg Trp Met Gly Pro Gln Arg Leu Met Phe Leu Val Leu Trp
    690                 695                 700

Lys Leu Ala Arg Gly Ala Phe Pro Leu Ala Leu Leu Met Gly Ile Ser
705                 710                 715                 720

Ala Thr Arg Gly Arg Thr Ser Val Leu Gly Ala Glu Phe Cys Phe Asp
                725                 730                 735

Val Thr Phe Glu Val Asp Thr Ser Val Leu Gly Trp Val Val Ala Ser
            740                 745                 750

Val Val Ala Trp Ala Ile Ala Leu Leu Ser Ser Met Ser Ala Gly Gly
        755                 760                 765

Trp Lys His Lys Ala Val Ile Tyr Arg Thr Trp Cys Lys Gly Tyr Gln
    770                 775                 780
```

```
Ala Val Arg Gln Arg Val Val Arg Ser Pro Leu Gly Glu Gly Arg Pro
785                 790                 795                 800

Thr Lys Leu Leu Thr Phe Ala Trp Cys Leu Ala Ser Tyr Ile Trp Pro
                805                 810                 815

Asp Ala Val Met Met Val Val Val Ala Leu Val Leu Leu Phe Gly Leu
            820                 825                 830

Phe Asp Ala Leu Asp Trp Ala Leu Glu Glu Leu Leu Val Ser Arg Pro
        835                 840                 845

Ser Leu Arg Arg Leu Ala Arg Val Val Glu Cys Cys Val Met Ala Gly
    850                 855                 860

Glu Lys Ala Thr Thr Ile Arg Leu Val Ser Lys Met Cys Ala Arg Gly
865                 870                 875                 880

Ala Tyr Leu Phe Asp His Met Gly Ser Phe Ser Arg Ala Val Lys Glu
                885                 890                 895

Arg Leu Leu Glu Trp Asp Ala Ala Leu Glu Pro Leu Ser Phe Thr Arg
            900                 905                 910

Thr Asp Cys Arg Ile Ile Arg Asp Ala Ala Arg Thr Leu Ser Cys Gly
        915                 920                 925

Gln Cys Val Met Gly Leu Pro Val Val Ala Arg Arg Gly Asp Glu Val
    930                 935                 940

Leu Ile Gly Val Phe Gln Asp Val Asn His Leu Pro Pro Gly Phe Val
945                 950                 955                 960

Pro Thr Ala Pro Val Val Ile Arg Arg Cys Gly Lys Gly Phe Leu Gly
                965                 970                 975

Val Thr Lys Ala Ala Leu Thr Gly Arg Asp Pro Asp Leu His Pro Gly
            980                 985                 990

Asn Val Met Val Leu Gly Thr Ala Thr Ser Arg Ser Met Gly Thr Cys
        995                1000                1005

Leu Asn Gly Leu Leu Phe Thr Thr Phe His Gly Ala Ser Ser Arg Thr
   1010                1015                1020

Ile Ala Thr Pro Val Gly Ala Leu Asn Pro Arg Trp Trp Ser Ala Ser
1025                1030                1035                1040

Asp Asp Val Thr Val Tyr Pro Leu Pro Asp Gly Ala Thr Ser Leu Thr
               1045                1050                1055

Pro Cys Thr Cys Gln Ala Glu Ser Cys Trp Val Ile Arg Ser Asp Gly
           1060                1065                1070

Ala Leu Cys His Gly Leu Ser Lys Gly Asp Lys Val Glu Leu Asp Val
       1075                1080                1085
```

```
Ala Met Glu Val Ser Asp Phe Arg Gly Ser Ser Gly Ser Pro Val Leu
   1090                1095                1100

Cys Asp Lys Gly His Ala Val Arg Met Leu Val Ser Val Leu His Ser
1105                1110                1115                1120

Gly Gly Arg Val Thr Ala Ala Arg Phe Thr Arg Pro Trp Thr Gln Val
                1125                1130                1135

Pro Thr Asp Ala Lys Thr Thr Thr Glu Pro Pro Pro Val Pro Ala Lys
            1140                1145                1150

Gly Val Phe Lys Glu Ala Pro Leu Phe Met Pro Thr Gly Ala Gly Lys
        1155                1160                1165

Ser Thr Arg Val Pro Leu Glu Tyr Gly Asn Met Gly His Lys Val Leu
    1170                1175                1180

Ile Leu Asn Pro Ser Val Ala Thr Val Arg Ala Met Gly Pro Tyr Met
1185                1190                1195                1200

Glu Arg Leu Ala Gly Lys His Pro Ser Ile Tyr Cys Gly His Asp Thr
                1205                1210                1215

Thr Ala Phe Thr Arg Ile Thr Asp Ser Pro Leu Thr Tyr Ser Thr Tyr
            1220                1225                1230

Gly Arg Phe Leu Ala Asn Pro Arg Gln Met Leu Arg Gly Val Ser Val
        1235                1240                1245

Val Ile Cys Asp Glu Cys His Ser His Asp Ser Thr Val Leu Leu Gly
    1250                1255                1260

Ile Gly Arg Val Arg Glu Leu Ala Arg Gly Cys Gly Val Gln Leu Val
1265                1270                1275                1280

Leu Tyr Ala Thr Ala Thr Pro Pro Gly Ser Pro Met Thr Gln His Pro
                1285                1290                1295

Ser Ile Ile Glu Thr Lys Leu Asp Val Gly Glu Ile Pro Phe Tyr Gly
            1300                1305                1310

His Gly Ile Pro Leu Glu Arg Met Arg Thr Gly Arg His Leu Val Phe
        1315                1320                1325

Cys His Ser Lys Ala Glu Cys Glu Arg Leu Ala Gly Gln Phe Ser Ala
    1330                1335                1340

Arg Gly Val Asn Ala Ile Ala Tyr Tyr Arg Gly Lys Asp Ser Ser Ile
1345                1350                1355                1360

Ile Lys Asp Gly Asp Leu Val Val Cys Ala Thr Asp Ala Leu Ser Thr
                1365                1370                1375

Gly Tyr Thr Gly Asn Phe Asp Ser Val Thr Asp Cys Gly Leu Val Val
            1380                1385                1390
```

Glu Glu Val Val Glu Val Thr Leu Asp Pro Thr Ile Thr Ile Ser Leu
1395 1400 1405

Arg Thr Val Pro Ala Ser Ala Glu Leu Ser Met Gln Arg Arg Gly Arg
1410 1415 1420

Thr Gly Arg Gly Arg Ser Gly Arg Tyr Tyr Tyr Ala Gly Val Gly Lys
1425 1430 1435 1440

Ala Pro Ala Gly Val Val Arg Ser Gly Pro Val Trp Ser Ala Val Glu
1445 1450 1455

Ala Gly Val Thr Trp Tyr Gly Met Glu Pro Asp Leu Thr Ala Asn Leu
1460 1465 1470

Leu Arg Leu Tyr Asp Asn Cys Pro Tyr Thr Ala Ala Val Ala Ala Asp
1475 1480 1485

Ile Gly Glu Ala Ala Val Phe Phe Ser Gly Leu Ala Pro Leu Arg Met
1490 1495 1500

His Pro Asp Val Ser Trp Ala Lys Val Arg Gly Val Asn Trp Pro Phe
1505 1510 1515 1520

Leu Val Gly Val Gln Arg Thr Met Cys Arg Glu Thr Leu Ser Pro Gly
1525 1530 1535

Pro Ser Asp Asp Pro Gln Trp Ala Gly Leu Lys Gly Pro Asn Pro Val
1540 1545 1550

Pro Leu Leu Leu Arg Trp Gly Asn Asp Leu Pro Ser Lys Val Ala Gly
1555 1560 1565

His His Ile Val Asp Asp Leu Val Arg Arg Leu Gly Val Ala Glu Gly
1570 1575 1580

Tyr Val Arg Cys Asp Ala Gly Pro Ile Leu Met Val Gly Leu Ala Ile
1585 1590 1595 1600

Ala Gly Gly Met Ile Tyr Ala Ser Tyr Thr Gly Ser Leu Val Val Val
1605 1610 1615

Thr Asp Trp Asp Val Lys Gly Gly Gly Ser Pro Leu Tyr Arg His Gly
1620 1625 1630

Asp Gln Ala Thr Pro Gln Pro Val Val Gln Val Pro Pro Val Asp His
1635 1640 1645

Arg Pro Gly Gly Glu Ser Ala Pro Ser Asp Ala Asn Thr Val Thr Asp
1650 1655 1660

Ala Val Ala Ala Ile Gln Val Asp Cys Asp Trp Ser Val Met Thr Leu
1665 1670 1675 1680

Ser Ile Gly Glu Val Leu Ser Leu Ala Gln Ala Lys Thr Ala Glu Ala
1685 1690 1695

Tyr Ala Ala Thr Thr Lys Trp Leu Ala Gly Cys Tyr Thr Gly Thr Arg
1700 1705 1710

Ala Val Pro Thr Val Ser Ile Val Asp Lys Leu Phe Ala Gly Gly Trp
1715 1720 1725

Ala Ala Val Val Gly His Cys His Ser Val Ile Ala Ala Ala Val Ala
1730 1735 1740

Ala Tyr Gly Ala Ser Arg Ser Pro Pro Leu Ala Ala Ala Ala Ser Tyr
1745 1750 1755 1760

Leu Met Gly Leu Gly Val Gly Gly Asn Ala Gln Thr Arg Leu Ala Ser
1765 1770 1775

Ala Leu Leu Leu Gly Ala Ala Gly Thr Ala Leu Gly Thr Pro Val Val
1780 1785 1790

Gly Leu Thr Met Ala Gly Ala Phe Met Gly Ser Ala Ser Val Ser Pro
1795 1800 1805

Ser Leu Val Thr Ile Leu Leu Gly Ala Val Gly Gly Trp Glu Gly Val
1810 1815 1820

Val Asn Ala Ala Ser Leu Val Phe Asp Phe Met Ala Gly Lys Leu Ser
1825 1830 1835 1840

Ser Glu Asp Leu Trp Tyr Ala Ile Pro Val Leu Thr Ser Pro Gly Ala
1845 1850 1855

Gly Leu Ala Gly Ile Ala Leu Gly Leu Val Leu Tyr Ser Ala Asn Asn
1860 1865 1870

Ser Gly Thr Thr Thr Trp Leu Asn Arg Leu Leu Thr Thr Leu Pro Arg
1875 1880 1885

Ser Ser Cys Ile Pro Asp Ser Tyr Phe Gln Gln Ala Asp Tyr Cys Asp
1890 1895 1900

Lys Val Ser Ala Val Leu Arg Arg Leu Ser Leu Thr Arg Thr Val Val
1905 1910 1915 1920

Ala Leu Val Asn Arg Glu Pro Lys Val Asp Glu Val Gln Val Gly Tyr
1925 1930 1935

Val Trp Asp Leu Trp Glu Trp Ile Met Arg Gln Val Arg Met Val Met
1940 1945 1950

Ala Arg Leu Arg Ala Leu Cys Pro Val Val Ser Leu Pro Leu Trp His
1955 1960 1965

Cys Gly Glu Gly Trp Ser Gly Glu Trp Leu Leu Asp Gly His Val Glu
1970 1975 1980

Ser Arg Cys Leu Cys Gly Cys Val Ile Thr Gly Asp Val Leu Asn Gly
1985 1990 1995 2000

```
Gln Leu Lys Asp Pro Val Tyr Ser Thr Lys Leu Cys Arg His Tyr Trp
               2005                2010                2015

Met Gly Thr Val Pro Val Asn Met Leu Gly Tyr Gly Glu Thr Ser Pro
           2020                2025                2030

Leu Leu Ala Ser Asp Thr Pro Lys Val Val Pro Phe Gly Thr Ser Gly
       2035                2040                2045

Trp Ala Glu Val Val Val Thr Pro Thr His Val Val Ile Arg Arg Thr
   2050                2055                2060

Ser Ala Tyr Lys Leu Leu Arg Gln Gln Ile Leu Ser Ala Ala Val Ala
2065                2070                2075                2080

Glu Pro Tyr Tyr Val Asp Gly Ile Pro Val Ser Trp Asp Ala Asp Ala
               2085                2090                2095

Arg Ala Pro Ala Met Val Tyr Gly Pro Gly Gln Ser Val Thr Ile Asp
           2100                2105                2110

Gly Glu Arg Tyr Thr Leu Pro His Gln Leu Arg Leu Arg Asn Val Ala
       2115                2120                2125

Pro Ser Glu Val Ser Ser Glu Val Ser Ile Asp Ile Gly Thr Glu Thr
   2130                2135                2140

Glu Asp Ser Glu Leu Thr Glu Ala Asp Leu Pro Pro Ala Ala Ala Ala
2145                2150                2155                2160

Leu Gln Ala Ile Glu Asn Ala Ala Arg Ile Leu Glu Pro His Ile Asp
               2165                2170                2175

Val Ile Met Glu Asp Cys Ser Thr Pro Ser Leu Cys Gly Ser Ser Arg
           2180                2185                2190

Glu Met Pro Val Trp Gly Glu Asp Ile Pro Arg Thr Pro Ser Pro Ala
       2195                2200                2205

Leu Ile Ser Val Thr Glu Ser Ser Pro Asp Glu Lys Thr Pro Ser Val
   2210                2215                2220

Ser Ser Ser Gln Glu Asp Thr Pro Ser Ser Asp Ser Phe Glu Val Ile
2225                2230                2235                2240

Gln Glu Ser Glu Thr Ala Glu Gly Glu Glu Ser Val Phe Asn Val Ala
               2245                2250                2255

Leu Ser Val Leu Lys Ala Leu Phe Pro Gln Ser Asp Ala Thr Arg Lys
           2260                2265                2270

Leu Thr Val Lys Met Ser Cys Cys Val Glu Lys Ser Val Thr Arg Phe
       2275                2280                2285

Phe Ser Leu Gly Leu Thr Val Ala Asp Val Ala Ser Leu Cys Glu Met
   2290                2295                2300
```

```
Glu Ile Gln Asn His Thr Ala Tyr Cys Asp Lys Val Arg Thr Pro Leu
2305                2310                2315                2320

Glu Leu Gln Val Gly Cys Leu Val Gly Asn Glu Leu Thr Phe Glu Cys
                2325                2330                2335

Asp Lys Cys Glu Ala Arg Gln Glu Thr Leu Ala Ser Phe Ser Tyr Ile
            2340                2345                2350

Trp Ser Gly Val Pro Leu Thr Arg Ala Thr Pro Ala Lys Pro Pro Val
        2355                2360                2365

Val Arg Pro Val Gly Ser Leu Leu Val Ala Asp Thr Thr Lys Val Tyr
    2370                2375                2380

Val Thr Asn Pro Asp Asn Val Gly Arg Arg Val Asp Lys Val Thr Phe
2385                2390                2395                2400

Trp Arg Ala Pro Arg Val His Asp Lys Phe Leu Val Asp Ser Ile Glu
                2405                2410                2415

Arg Ala Lys Arg Ala Ala Gln Ala Cys Leu Ser Met Gly Tyr Thr Tyr
            2420                2425                2430

Glu Glu Ala Ile Arg Thr Val Arg Pro His Ala Ala Met Gly Trp Gly
        2435                2440                2445

Ser Lys Val Ser Val Lys Asp Leu Ala Thr Pro Ala Gly Lys Met Ala
    2450                2455                2460

Val His Asp Arg Leu Gln Glu Ile Leu Glu Gly Thr Pro Val Pro Phe
2465                2470                2475                2480

Thr Leu Thr Val Lys Lys Glu Val Phe Phe Lys Asp Arg Lys Glu Glu
                2485                2490                2495

Lys Ala Pro Arg Leu Ile Val Phe Pro Pro Leu Asp Phe Arg Ile Ala
            2500                2505                2510

Glu Lys Leu Ile Leu Gly Asp Pro Gly Arg Val Ala Lys Ala Val Leu
        2515                2520                2525

Gly Gly Ala Tyr Ala Phe Gln Tyr Thr Pro Asn Gln Arg Ile Arg Glu
    2530                2535                2540

Met Leu Lys Leu Trp Glu Ser Lys Lys Thr Pro Cys Ala Ile Cys Val
2545                2550                2555                2560

Asp Ala Thr Cys Phe Asp Ser Ser Ile Thr Glu Glu Asp Val Ala Leu
                2565                2570                2575

Glu Thr Glu Leu Tyr Ala Leu Ala Ser Asp His Pro Glu Trp Val Arg
            2580                2585                2590

Ala Leu Gly Lys Tyr Tyr Ala Ser Gly Thr Met Val Thr Pro Glu Gly
        2595                2600                2605
```

Val Pro Val Gly Glu Arg Tyr Cys Arg Ser Ser Gly Val Leu Thr Thr
2610 2615 2620

Ser Ala Ser Asn Cys Leu Thr Cys Tyr Ile Lys Val Lys Ala Ala Cys
2625 2630 2635 2640

Glu Arg Val Gly Leu Lys Asn Val Ser Leu Leu Ile Ala Gly Asp Asp
2645 2650 2655

Cys Leu Ile Ile Cys Glu Arg Pro Val Cys Asp Pro Ser Asp Ala Leu
2660 2665 2670

Gly Arg Ala Leu Ala Ser Tyr Gly Tyr Ala Cys Glu Pro Ser Tyr His
2675 2680 2685

Ala Ser Leu Asp Thr Ala Pro Phe Cys Ser Thr Trp Leu Ala Glu Cys
2690 2695 2700

Asn Ala Asp Gly Lys Arg His Phe Phe Leu Thr Thr Asp Phe Arg Arg
2705 2710 2715 2720

Pro Leu Ala Arg Met Ser Ser Glu Tyr Ser Asp Pro Met Ala Ser Ala
2725 2730 2735

Ile Gly Tyr Ile Leu Leu Tyr Pro Trp His Pro Ile Thr Arg Trp Val
2740 2745 2750

Ile Ile Pro His Val Leu Thr Cys Ala Phe Arg Gly Gly Gly Thr Pro
2755 2760 2765

Ser Asp Pro Val Trp Cys Gln Val His Gly Asn Tyr Tyr Lys Phe Pro
2770 2775 2780

Leu Asp Lys Leu Pro Asn Ile Ile Val Ala Leu His Gly Pro Ala Ala
2785 2790 2795 2800

Leu Arg Val Thr Ala Asp Thr Thr Lys Thr Lys Met Glu Ala Gly Lys
2805 2810 2815

Val Leu Ser Asp Leu Lys Leu Pro Gly Leu Ala Val His Arg Lys Lys
2820 2825 2830

Ala Gly Ala Leu Arg Thr Arg Met Leu Arg Ser Arg Gly Trp Ala Glu
2835 2840 2845

Leu Ala Arg Gly Leu Leu Trp Arg Pro Gly Leu Arg Leu Pro Pro Pro
2850 2855 2860

Glu Ile Ala Gly Ile Pro Gly Gly Phe Pro Leu Ser Pro Pro Tyr Met
2865 2870 2875 2880

Gly Val Val His Gln Leu Asp Phe Thr Ser Gln Arg Ser Arg Trp Arg
2885 2890 2895

Trp Leu Gly Phe Leu Ala Leu Leu Ile Val Ala Leu Phe Gly
2900 2905 2910